

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

日本国特許庁
JAPAN PATENT OFFICE

T. Miyakawa et al.
12/8/03
Q 78853
18f1

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application: 2002年12月10日

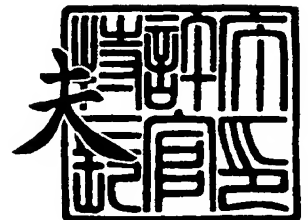
出願番号
Application Number: 特願2002-358407
[ST. 10/C]: [JP 2002-358407]

出願人
Applicant(s): 日本電気株式会社

2003年10月 2日

特許庁長官
Commissioner,
Japan Patent Office

今井 康夫



出証番号 出証特2003-3081302

【書類名】 特許願

【整理番号】 64002111

【提出日】 平成14年12月10日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 17/30

【発明者】

 【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

 【氏名】 宮川 知也

【発明者】

 【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

 【氏名】 仲里 猛留

【発明者】

 【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

 【氏名】 麻生川 稔

【発明者】

 【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

 【氏名】 剣持 聡久

【特許出願人】

 【識別番号】 000004237

 【氏名又は名称】 日本電気株式会社

【代理人】

 【識別番号】 100088328

 【弁理士】

 【氏名又は名称】 金田 暢之

 【電話番号】 03-3585-1882

【選任した代理人】

 【識別番号】 100106297

 【弁理士】

 【氏名又は名称】 伊藤 克博

【選任した代理人】

【識別番号】 100106138

【弁理士】

【氏名又は名称】 石橋 政幸

【手数料の表示】

【予納台帳番号】 089681

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9710078

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 配列表示方法／装置／プログラム／記録媒体、相同性検索方法／装置／プログラム／記録媒体

【特許請求の範囲】

【請求項 1】 類似した複数の塩基配列またはアミノ酸配列の表示方法において、

前記類似した複数の塩基配列またはアミノ酸配列における変異性および／または共通性に応じた視覚的特徴を付加することを特徴とする配列表示方法。

【請求項 2】 前記視覚的特徴を前記類似した複数の塩基配列またはアミノ酸配列における変異部位および／または共通部位に付加する、請求項 1 に記載の配列表示方法。

【請求項 3】 前記視覚的特徴を前記類似した複数の塩基配列またはアミノ酸配列における変異性および／または共通性の程度に応じて付加する、請求項 1 または 2 に記載の配列表示方法。

【請求項 4】 前記視覚的特徴を前記類似した複数の塩基配列またはアミノ酸配列の前記変異部位における変異の頻度に応じて付加する、請求項 2 または 3 に記載の配列表示方法。

【請求項 5】 前記視覚的特徴を、前記類似した複数の塩基配列におけるコドンに対応する、アミノ酸名および性質を含むアミノ酸情報に基づいて付加する、請求項 2 から 4 のいずれか 1 項に記載の配列表示方法。

【請求項 6】 前記視覚的特徴を、表示色を用いて表現する、請求項 1 から 5 のいずれか 1 項に記載の配列表示方法。

【請求項 7】 前記視覚的特徴を、表示色の変化を用いて表現する、請求項 1 から 6 のいずれか 1 項に記載の配列表示方法。

【請求項 8】 前記視覚的特徴を、文字の種類を用いて表現する、請求項 1 から 5 のいずれか 1 項に記載の配列表示方法。

【請求項 9】 前記類似した複数の塩基配列またはアミノ酸配列を、各配列中の対応する塩基またはアミノ酸を揃えて表示する、請求項 1 から 8 のいずれか 1 項に記載の配列表示方法。

【請求項 10】 前記類似した複数の塩基配列またはアミノ酸配列どうしのリンクおよび／または関連する情報へのリンクを付加する、請求項 1 から 9 のいずれか 1 項に記載の配列表示方法。

【請求項 11】 類似した複数の塩基配列またはアミノ酸配列の表示装置において、

前記類似した複数の塩基配列またはアミノ酸配列における変異性および／または共通性に応じた視覚的特徴を付加する手段を有することを特徴とする配列表示装置。

【請求項 12】 前記手段は、前記視覚的特徴を前記類似した複数の塩基配列またはアミノ酸配列における変異部位および／または共通部位に付加する、請求項 11 に記載の配列表示装置。

【請求項 13】 前記手段は、前記視覚的特徴を前記類似した複数の塩基配列またはアミノ酸配列における変異性および／または共通性の程度に応じて付加する、請求項 11 または 12 に記載の配列表示装置。

【請求項 14】 前記手段は、前記視覚的特徴を前記類似した複数の塩基配列またはアミノ酸配列の前記変異部位における変異の頻度に応じて付加する、請求項 12 または 13 に記載の配列表示装置。

【請求項 15】 前記手段は、前記視覚的特徴を、前記類似した複数の塩基配列におけるコドンに対応する、アミノ酸名および性質を含むアミノ酸情報に基づいて付加する、請求項 12 から 14 のいずれか 1 項に記載の配列表示装置。

【請求項 16】 前記手段は、前記視覚的特徴を、表示色を用いて表現する、請求項 11 から 15 のいずれか 1 項に記載の配列表示装置。

【請求項 17】 前記手段は、前記視覚的特徴を、表示色の変化を用いて表現する、請求項 11 から 16 のいずれか 1 項に記載の配列表示装置。

【請求項 18】 前記手段は、前記視覚的特徴を、文字の種類を用いて表現する、請求項 11 から 15 のいずれか 1 項に記載の配列表示装置。

【請求項 19】 前記類似した複数の塩基配列またはアミノ酸配列を、各配列中の対応する塩基またはアミノ酸を揃えて表示する手段をさらに有する、請求項 11 から 18 のいずれか 1 項に記載の配列表示装置。

【請求項 2 0】 前記類似した複数の塩基配列またはアミノ酸配列どうしのリンクおよび／または関連する情報へのリンクを付加する手段をさらに有する、請求項 1 1 から 1 9 のいずれか 1 項に記載の配列表示装置。

【請求項 2 1】 請求項 1 から 1 0 のいずれか 1 項に記載の配列表示方法をコンピュータに実行させる配列表示プログラム。

【請求項 2 2】 請求項 1 から 1 0 のいずれか 1 項に記載の配列表示方法をコンピュータに実行させる配列表示プログラムを記録した、コンピュータで読み取りが可能な記録媒体。

【請求項 2 3】 ユーザによって入力された生物情報データベースへの問い合わせを解析するステップと、

該問い合わせの解析をもとに、前記生物情報データベースに適合する検索条件を生成して該生物情報データベースを検索するステップと、

該検索結果を解析するステップと、

該解析結果と請求項 1 から 1 0 のいずれか 1 項に記載の配列表示方法を用いて、前記検索結果を表示するステップを有する相同性検索方法。

【請求項 2 4】 ユーザによって入力された生物情報データベースへの問い合わせを解析するクエリ解析手段と、

該問い合わせの解析をもとに、前記生物情報データベースに適合する検索条件を生成して該生物情報データベースを検索する通信手段と、

該検索結果を解析する結果解析手段と、

該解析結果と請求項 1 1 から 2 0 のいずれか 1 項に記載の配列表示装置を用いて、前記検索結果を表示する表示手段を有する相同性検索装置。

【請求項 2 5】 前記クエリ解析手段と前記通信手段と前記結果解析手段と前記表示手段の動作を制御する制御手段をさらに有する、請求項 2 4 に記載の相同性検索装置。

【請求項 2 6】 請求項 2 3 に記載の相同性検索方法をコンピュータに実行させる相同性検索プログラム。

【請求項 2 7】 請求項 2 3 に記載の相同性検索方法をコンピュータに実行させる相同性検索プログラムを記録した、コンピュータで読み取りが可能な記録

媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、塩基配列またはアミノ酸配列の相同性検索に関する。

【0002】

【従来の技術】

現在、ヒトゲノム配列（ヒトのすべてのDNA配列）がほぼ解読され、アメリカ国立健康研究所NIH（National Institutes of Health）に所属する一機関であるアメリカ国立医学図書館NLM（National Library of Medicine）の生物工学情報センターNCBI（National Center for Biotechnology Information）、日本の国立遺伝学研究所、欧州の生物情報研究機関であるEBI（European Bioinformatics Institute）およびヨーロッパ分子生物学研究所EMBL（European Molecular Biology Laboratory）、日本の東京大学医科学研究所等によって、DNA配列およびDNA配列に由来するアミノ酸配列などがデータベース（DB）として構築されている。

【0003】

他の生物種についても、マウス、ラット、フグ、ゼブラフィッシュ、ショウジョウバエ、線虫などのゲノムが精力的に解読されつつあるか、すでに解読が完了している。

【0004】

さらに次の段階として、ゲノム配列における個人差を調べる試みも進みつつある。

【0005】

このような状況において、生物学および医学の分野では、研究者がその研究の過程で得た塩基配列またはアミノ酸配列と既知の配列との相同性比較を行うことが非常に重要である。なぜなら、他の研究者によって相同性の非常に高い配列が

すでに登録されている場合、その配列に由来する生体機能がある程度解析済みであることを意味し、自身の研究の方向性を変更しなければならないからである。逆に、まだ登録されていない場合には、既知の配列から機能を予測することも必要となる。

【0006】

BLAST (Basic Local Alignment Search Tool、<http://www.ncbi.nlm.nih.gov/blast/>) は、NCBIが開発した塩基配列およびアミノ酸配列の相同性検索システムである（例えば、非特許文献1参照）。これを用いると、NCBIが提供する塩基配列およびアミノ酸配列をはじめとするDBであるEntrez (<http://www.ncbi.nlm.nih.gov/entrez/>) などから、相同性の高い配列情報を取得することができる。このため、全世界の生物医学系の研究者により日常的に幅広く利用されている。

【0007】

BLAST以外の塩基配列およびアミノ酸配列の相同性検索システムとしては、FASTA (Fast Alignment) がよく利用されている（例えば、非特許文献2参照）。配列DBに関しても、NCBIのDB以外に、日本の国立遺伝学研究所が運営する日本DNAデータバンクDDBJ (DNA Data Bank of Japan)、EBI/EMBLが運営する塩基配列およびアミノ酸配列DBなどがある。さらに、日本の東京大学医科学研究所のヒトの一塩基置換DBであるJ SNPsなど、その目的に応じて様々な塩基配列およびアミノ酸配列DBが構築され、各研究者の目的に応じて利用されている。

【0008】

上記各種の配列情報DBに対して、塩基配列またはアミノ酸配列の相同性検索を行う場合、研究者は主に上記のBLASTまたはFASTAシステムを用いることが多い。

【0009】

さらに、相同性検索を用いた配列解析では、3つ以上の塩基配列またはアミノ酸配列について、相同性に基づいて対応する部分をそろえて並べる機能（以下、

多重配列比較またはマルチプルアライメントと記す) が必要なことも多い。マルチプルアライメントを実行するソフトウェアとしては例えば、EBI/EMBLで開発されたClustal W、Clustal Xなどがあり、その他にも日本国内において開発されたソフトウェアを含めて数多くの多重配列比較ソフトウェアが存在し、研究者に利用されている。

【0010】

【非特許文献1】

ジャーナル・オブ・モリキュラー・バイオロジー、215 (1990) 403~410ページ (Journal of Molecular Biology, 215 (1990) pp403~410)

【非特許文献2】

ドゥーリトル (Doolittle) 編、メソッド・イン・エンザイモロジー、183 (1990) 63~98ページ (Methods in Enzymology, 183 (1990) pp63~98)、アカデミックプレス (Academic Press)

【0011】

【発明が解決しようとする課題】

塩基配列またはアミノ酸配列の相同性検索において、ユーザが入力した塩基配列またはアミノ酸配列、あるいは、ユーザが入力した情報により導かれた塩基配列またはアミノ酸配列が、DBへの問い合わせ（以下、クエリと呼ぶ）となり、類似の配列が検索される。検索結果においては、相同性により対応付けられる部位で、クエリとした塩基配列またはアミノ酸配列と、検索結果の塩基配列またはアミノ酸配列との間で、塩基またはアミノ酸が異なる（以下、変異と呼ぶ）箇所および同じである（以下、共通と呼ぶ）箇所が重要な意味を持っている。BLASTまたはFASTAシステムを用いる場合、以下のような問題点がある。

【0012】

1. クエリとした配列と検索された配列が対で表示されるため、多数の配列間における変異の部位および頻度などの重要な意味を持つ情報を得るには、ユーザがすべての対の比較結果を相互参照しなければならない。

【0013】

2. 検索の結果は、相同性の高い部位だけが切り取られて部分的に表示されるので、全体の配列からみてどの部分が表示されているのかが非常にわかりにくい。

【0014】

3. 比較した部位の塩基またはアミノ酸が同じ場合は「|」や「.」で表され、結果を一目で把握することが困難である。

【0015】

4. 検索結果が単純なテキスト形式でのみ表示されるため、変異部位を見落としやすい。

【0016】

以上まとめると、BLASTまたはFASTAシステムの検索結果表示では、ユーザが最も必要とする、配列間の変位部位または共通部位などの重要な情報へのアクセスが容易ではない点が問題となる。

【0017】

上記問題点に対処するために、研究者はしばしば多重配列比較（マルチプルアライメント）を行う場合がある。しかしながら、多重配列比較を行う上記のソフトウェアを用いても、上記問題点は十分に解決されない。

【0018】

すなわち、例えば上記のClustal Xは、アミノ酸の変異部位を色付き表示する機能もあるが、色付けの基準が一定ではなく、変異があることを強調する以上の意味を有していない。

【0019】

さらに、用いるソフトウェアによっては相同性の求め方がBLASTやFASTAのそれとは異っており、BLASTやFASTAの検索結果に対して多重配列比較を行うと、しばしば塩基位置の対応がずれる。このような場合、例えばBLASTで得た検索結果を基準として塩基位置の対応を合わせる手作業が必要となる。

【0020】

加えて、これらのソフトウェアを用いて相同性を求めるには、比較を行う全ての配列情報を予め収集しておく必要があり、収集すべき配列情報が多量にある場合、収集漏れが生じたり、漏れがないか検証する作業が研究者への大きな負担となっていたりした。

【0021】

そこで本発明の目的は、配列間の変位部位や共通部位などに関する情報へのアクセスが容易で、なおかつ、検索された複数の配列がそのまま結果表示される配列表示方法および相同性検索方法を提供する。

【0022】

【課題を解決するための手段】

上記目的を達成するために、本発明の配列表示方法は、類似した複数の塩基配列またはアミノ酸配列における変異性および／または共通性に応じた視覚的特徴を付加することを特徴としている。

【0023】

視覚的特徴は、類似した複数の塩基配列またはアミノ酸配列における変異部位および／または共通部位に付加される。

【0024】

さらに、視覚的特徴は、類似した複数の塩基配列またはアミノ酸配列における変異性および／または共通性の程度に応じて付加される。

【0025】

さらに、視覚的特徴は、類似した複数の塩基配列またはアミノ酸配列の変異部位における変異の頻度に応じて付加される。

【0026】

以上により、ユーザは、ユーザ自身による相互参照なしに、変異性および共通性などの重要な情報を即座に得ることができる。

【0027】

さらに、視覚的特徴は、塩基配列をアミノ酸配列に変換して、塩基配列における変異部位のコドンに対応する、アミノ酸名および性質を含むアミノ酸情報に基づいて付加される。これにより、塩基配列の変異がアミノ酸生成に及ぼす影響を

評価することが可能になる。

【0028】

これらの視覚的特徴は、表示色、さらに表示色の変化、または文字の種類を用いて表現される。

【0029】

また、類似した複数の塩基配列またはアミノ酸配列は、各配列中の対応する塩基またはアミノ酸を揃えて表示される。これにより、マルチプルアライメントの結果においても、塩基またはアミノ酸の変異および共通に関する部位、程度、頻度等をユーザが容易に認識することができる。

【0030】

また、類似した複数の塩基配列またはアミノ酸配列に関する情報へのリンクを設ける。これにより、相同性検索結果が含む情報どうし、および、相同性検索結果が含む情報とインターネット上の情報のような外部の情報との関連付けも、表示情報が含むことができるようになり、相互参照や参照情報の閲覧も容易に行えるようになる。

【0031】

本発明の相同性検索方法は、ユーザによって入力された生物情報データベースへの問い合わせを解析するステップと、この問い合わせの解析をもとに、生物情報データベースに適合する検索条件を生成して生物情報データベースを検索するステップと、検索結果を解析するステップと、解析結果と上述の配列表示方法を用いて、検索結果を表示するステップからなる。以上により、検索から表示が一体に行え、多重配列比較を行う場合にも、収集漏れが生じたり、漏れがないか検証したりする必要はない。

【0032】

また、相同性検索装置に制御部を設け、各上記ステップの処理を制御させるようにしてもよい。

【0033】

【発明の実施の形態】

本発明の実施の形態について、図面を参照して詳細に説明する。

【0034】

図1を参照すると、本発明の第1の実施形態の相同性検索システム10は、クエリ解析部11と、通信部12と、結果解析部13と、表示部14を有する。

【0035】

クエリ解析部11は、ユーザが入力した生物情報DBのクエリ情報（クエリ）を受け取って検索条件の解析を行い、生物情報DBの検索条件情報を通信部12に出力する。入力されるクエリは、塩基配列またはアミノ酸配列が主な情報として含まれる。他に、生物学的情報としてアクセッション番号、実験データ、ファイルなどを情報として含めることができる。加えて、DBの指定、結果を絞り込むための生物種、塩基配列ならばアミノ酸に翻訳するか、アミノ酸配列ならばP S I-B L A S Tを行うかの指定の情報を含めることができる。

【0036】

通信部12は、クエリ解析部11からの検索条件情報を受け取り、サーバ21を介して配列情報DB22に送り、配列情報DB22からサーバ21を介して検索結果を受け取って検索結果および検索条件情報を結果解析部13に出力する。

【0037】

結果解析部13は、通信部12から検索条件情報および検索結果を受け取り、検索結果の解析を行って、表示部14に解析結果を出力する。検索処理によって生成された結果を解析し、例えば変異の部位、頻度、程度といった情報を得る。

【0038】

表示部14は、結果解析部13から解析結果を受け、視覚的情報（例えば文字の色、文字の種類など）を付けた上で表示情報を生成し、出力する。結果解析部13によって機能づけされた情報を、視覚的特性（例えば文字の色）を付ける、またはリンクを生成するなど行い、適切な表示形式に変換し、表示を行う。

【0039】

図2を参照すると、クエリ解析部11における処理のフローチャートが示されている。

【0040】

ステップ201で、入力されたクエリに基づいて相同性検索および検索結果の

解析をどのように行うかを決定する。クエリが配列データまたはアクセッション番号等であった場合、相同性検索および検索結果の解析を行うようにフラグを設定する。クエリが検索結果ファイルを指定するものであった場合、相同性検索を行わないようにフラグを設定する。

【 0 0 4 1 】

ステップ 2 0 2 で、入力されたクエリに基づき、クエリが含む情報から配列情報への変換を行うかどうかのフラグを設定する。例えば、クエリがアクセッション番号を含み、ユーザが配列情報への変換を行うことを指定したような場合、ステップ 2 0 3 で、配列情報変換フラグの値を真にする。

【 0 0 4 2 】

ステップ 2 0 4 で、入力されたクエリに基づき、入力された、または入力された生物情報により導かれた配列が、塩基配列かアミノ酸配列かを判定する。

【 0 0 4 3 】

ステップ 2 0 4 で塩基配列と判定された場合、ステップ 2 0 5 で、ユーザが比較すべき塩基配列 DB として何を指定したかの情報を取得する。塩基配列 DB としては例えば、nr (Non-Redundant、標準となる非冗長配列 DB)、EST (Expressed Sequence Tag、発現している遺伝子の配列 DB)、SNPs (Single Nucleotide Polymorphisms、一塩基多型配列 DB) などである。

【 0 0 4 4 】

ステップ 2 0 6 で、入力されたクエリに基づき、入力されたクエリの、または入力されたクエリによって導かれる塩基配列をアミノ酸配列に変換して、さらなる解析を行うかどうかを判断する。変換する場合、ステップ 2 0 7 でアミノ酸変換フラグの値を真にする。

【 0 0 4 5 】

ステップ 2 0 4 で配列がアミノ酸配列だったと判定された場合、ステップ 2 0 8 で、ユーザが比較すべきアミノ酸配列 DB として何を指定したかの情報を取得する。アミノ酸配列 DB としては例えば、nr (non-Redundant、標準となる非冗長配列 DB)、PDB (Protein Data Bank、

タンパク質立体構造DB)、swissprot (アミノ酸配列DB)、patent (特許における配列DB)、yeast (酵母の配列DB)、month (直近に追加された配列DB) などである。

【0046】

ステップ209で、入力されたクエリに基づき、PSI-BLAST (Position Specific Interactive Basic Local Alignment Search Tool) を行うかどうかを判断する。行う場合、ステップ210でPSI-BLASTフラグの値を真にする。

【0047】

ステップ211で、入力されたクエリに基づき、検索結果を生物種によって絞り込むかどうかを判断する。ここで生物種とは例えば、ウイルス、バクテリア、菌類、真核生物、哺乳類、げっ歯類、シロイヌナズナ、枯草菌、線虫、ゼブラフィッシュ、キイロショウジョウバエ、大腸菌、ヒト、マウス、ラット、アフリカツメガエルなどである。ユーザが絞り込むと指定している場合、ステップ212で生物種情報フラグの値を真にする。

【0048】

図3を参照すると、通信部12における処理のフローチャートが示されている。

【0049】

ステップ301で、クエリ解析部11で解析されたクエリに関し、あらかじめ生物情報DBの情報を保持しておき、どの生物情報DBに問い合わせるかを決定する。もちろん、ユーザによってクエリを行う生物情報DBの情報を設定するようにすることも容易に実現できる。

【0050】

DBの場所は、大きく分ければ、ローカルのみ、ローカルとインターネットの両方、インターネットのみである。ここでローカルとは、本実施形態を実装するコンピュータ自身およびLAN (ローカルエリアネットワーク) で接続された範囲を指す。インターネットとは、WAN (ワイドエリアネットワーク、電話回線などを利用した接続) の範囲をさす。各々、複数のDBにクエリを振り分けるこ

とができる。

【 0 0 5 1 】

ローカルの場合、ステップ 3 0 2 で、クエリ解析部 1 1 で解析された塩基配列またはアミノ酸配列の情報を DB に見合った形式で作成する。

【 0 0 5 2 】

ステップ 3 0 3 で、ステップ 3 0 2 で作成されたクエリをローカルのサーバに渡し、相同性検索を実行させる。

【 0 0 5 3 】

ステップ 3 0 4 で、ステップ 3 0 3 で実行させた相同性検索の結果を統合する。これは、ローカルのみ、ローカルとインターネット、インターネットのみのそれぞれで、複数の DB にクエリを渡し、実行させるからである。

【 0 0 5 4 】

インターネットとローカルの両方の場合、ステップ 3 0 5 で、クエリ解析部 1 1 で解析された塩基配列またはアミノ酸配列の情報を DB に見合った形式で作成する。

【 0 0 5 5 】

ステップ 3 0 6 で、ステップ 3 0 5 で作成されたクエリをローカルのサーバに渡し、相同性検索を実行させる。ステップ 3 0 7 で、ステップ 3 0 5 で作成されたクエリをインターネット上のサーバに渡し、相同性検索を実行させる。

【 0 0 5 6 】

ステップ 3 0 8 で、ステップ 3 0 6 およびステップ 3 0 7 で実行させた相同性検索の結果を統合する。

【 0 0 5 7 】

インターネットの場合、ステップ 3 0 9 で、クエリ解析部 1 1 で解析された塩基配列またはアミノ酸配列の情報を DB に見合った形式で作成する。

【 0 0 5 8 】

ステップ 3 1 0 で、ステップ 3 0 9 で作成されたクエリをインターネット上のサーバに渡し、相同性検索を実行させる。

【 0 0 5 9 】

ステップ 311 で、ステップ 310 で実行させた相同性検索の結果を統合する。

【0060】

ステップ 312 で、ステップ 304 またはステップ 308 またはステップ 311 で統合された検索結果を保存する。

【0061】

図 4 を参照すると、結果解析部 13 における処理のフローチャートが示されている。

【0062】

ステップ 401 で、通信部 12 から受信した検索結果をとりこむ。もしも、相同性検索を行わない場合、ユーザによって指定されたファイルをとりこむ。

【0063】

ステップ 402 で、ステップ 401 でとりこんだ検索結果を解析し、クエリの配列での変異部位を同定し、その情報を取得する。

【0064】

ステップ 403 で、ステップ 402 で得られた変異部位の情報を統合し、その変異が比較した配列の中でどのくらいの頻度で起きているかを計算する。

【0065】

ステップ 404 で、ステップ 403 で得られた変異部位において、変異の程度を計算する。クエリが塩基配列で、図 2 のステップ 207 で塩基配列をアミノ酸に翻訳して解析するフラグが設定された場合、入力された塩基配列と変異による塩基配列をそれぞれアミノ酸配列に変換し、これらの比較を行う。

【0066】

ここでいう変異の程度とは例えば、塩基配列ならばアミノ酸置換を伴うか、または、アミノ酸配列ならば疎水性度や酸性／塩基性度などの性質がどの程度変わるのかをいう。アミノ酸配列またはアミノ酸に翻訳した塩基配列の場合、例えば、BLOSUM62 という変換行列により、その性質の近さを計算できる。ただし、これらはあくまでも例であって、酸性度、疎水性度、等電点によって計算するなど、これらに限定されるものではない。

【0067】

ステップ405で、表示部14に解析結果を出力する。

【0068】

図4を参照すると、表示部14における処理のフローチャートが示されている。

【0069】

ステップ501では、図4のステップ402によって解析された変異部位の情報をもとに、対応する部位をそろえるよう並べなおす。

【0070】

ステップ502では、図4のステップ402によって解析された変異部位の情報をもとに、変異部位の視覚的特性（例えば文字の色や文字の種類など）を決定する。

【0071】

ここで視覚的特性とは例えば、文字についてフォント、大文字／小文字、色など、地について色、パターン、テクスチャ、アニメーションなどがある。ただし、これらはあくまでも例であって、文字のサイズ、太字、下線、斜体、点滅、影つき、白抜き、箱囲み、地のテクスチャ、アニメーションなども含まれる。視覚的特性に関しても同様である。

【0072】

ステップ503で、図4のステップ403によって解析された変異の頻度の情報をもとに、変異の頻度の視覚的特性（例えば地の色）を決定する。ここでは、表示色の変化を用いて頻度を表現することが望ましい。ここでいう表示色の変化とは、例えば色の濃淡や色味の変化（一例として、赤から青へなど）などをさす。

【0073】

ステップ504で、図4のステップ404によって解析された変異の程度の情報をもとに、変異の程度の視覚的特性（例えば地の色）を決定する。ここでは、表示色の変化を用いて頻度を表現することが望ましい。ここでいう表示色の変化とは、例えば色の濃淡や色味の変化（一例として、赤から青へなど）などをさす。

【0074】

ステップ505で、相同性検索の結果が含む情報をリンクにより結ぶ。

【0075】

ステップ506で、図5のステップ501からステップ505によってつけられた視覚的特性の情報をもとに、出力に合わせて適切な表示形式に変換する。ここでいう出力とは例えば、画面やプリンタなどがあり、表示形式とは例えば、HTML (Hyper Text Markup Language)、XML (eXtensible Markup Language) などがある。ただし、これらはあくまでも例であって、TeX、bmp、gif、jpeg、PNG、TIFF、PICT、PDF (Portable Document Format)、PostScriptを含む。

【0076】

次に、第1の実施形態の具体的な実施例について、図面を参照して詳細に説明する。以下は、図2～図5のフローチャートをコンピュータプログラムとして作成して記録媒体に記録し、例えばパーソナルコンピュータに読み取らせて動作させる場合を示す。

【0077】

図6は、塩基配列について相同性検索を行う際の入力画面を示している。ここでは、ユーザによって、塩基配列の入力、検索を行うDBの指定、アミノ酸に翻訳して解析するかどうかの指定、検索結果を限定するための生物種の指定、まだ解析が行われていない検索結果ファイルの指定が行える。

【0078】

ユーザが塩基配列入力フォームに塩基配列を入力した後、Submitボタンをクリックすると、クエリ解析部11が塩基配列を受け取り、続いて上述のような各部の処理が行われる。

【0079】

本実施例では、ユーザが、アミノ酸に翻訳して解析するかどうかを指定できるようになっており、ここでの指定をアミノ酸変換フラグとして保持する。フラグ

が真ならば、アミノ酸変換手段により入力された塩基配列をアミノ酸配列に変換したうえで、変異の程度をアミノ酸配列も考慮に入れて計算する。これにより、ユーザは、その塩基の変異がアミノ酸置換を伴うかどうかを容易に把握し、実験操作によりその変異を元に戻さねばならないか否かの示唆が得られる。

【0080】

また、本実施例では、比較するDBをユーザが指定できるようになっており、これにより、あらかじめ保持してある生物情報DBの情報をもとに、クエリを振り分ける。例えば、ユーザがクエリとしてある塩基配列、および生物情報DBとしてnrおよびESTを指定すると、この塩基配列をnrとEST、それぞれで相同性検索を行うように検索条件情報を通信部12に送る。これにより、ユーザは、いくつかのDBを指定することで必要な情報を網羅的に得られる一方、用途に合わせたDBを選ぶことで冗長な検索結果を排除することができ、検索結果を容易に把握することができる。

【0081】

また、本実施例では、検索結果を生物種によって絞り込むかどうかをユーザが指定することができるようになっている。ここでの指定を生物種情報フラグに保持する。フラグが真ならば、特定の生物種の配列との比較結果のみを表示する。例えば、ユーザは生物種として、ヒト、マウス、ラットなどが指定できる。これにより、ユーザは、冗長な検索結果を排除することができ、検索結果を容易に把握することができる。

【0082】

また、本実施例では、ユーザが従来の相同性検索によって得られた結果を解析対象として選択することができるようになっている。例えば、結果がまだ未解析の解析結果ファイルをユーザが指定できる。これにより、相同性検索を行わず、結果解析のみを行うようにフラグを設定する。そのためユーザは、本発明によらない検索結果に対しても、本発明を用いた解析によりこの検索結果を把握することができる。

【0083】

図7は、塩基配列について相同性検索を行う際の結果画面の一部を示している

。入力された塩基配列のうち、変異のある部位は赤字で表している。塩基配列の下に、対応するアミノ酸配列が記載されており、特にアミノ酸置換を伴う変異部位は小文字のアルファベットで示されている。これにより、その塩基の変異が、アミノ酸置換を伴うかどうかをユーザは容易に把握することができ、実験操作により、その変異を元に戻さねばならないかどうかをユーザは判定できる。

【 0 0 8 4 】

また、各々の変異部位に、図 8 にて後述する詳細な配列比較結果へのリンクの情報がつけられている。これにより、ユーザは入力した配列の全体的な変異の状況と各変異の詳細な状況を相互参照することが可能である。

【 0 0 8 5 】

図 8 は、塩基配列について相同性検索を行う際の結果画面の別の一部を示している。各々の配列は 6 0 文字ごとに区切られている。文字数は、例えば 9 0 文字、1 2 0 文字、または表示できるだけの文字数であってもよい。本実施例では、塩基 3 つでアミノ酸 1 つに対応して翻訳されることを考慮して、3 で割り切れ、なおかつ、ユーザへの分かりやすさから 1 0 で割り切れる、すなわち 3 0 の倍数の文字数ごとに区切っている。

【 0 0 8 6 】

また、検索結果の中に、変異の中でも、特に塩基の挿入が見られた部位では、地の色を黄色にして特に強調している。これは、塩基の挿入または欠落は、その部位以降のアミノ酸配列がまったく変わってしまうことを意味し、研究者にとって、特に重要な情報だからである。

【 0 0 8 7 】

また、各々の変異において、それがどのくらいの頻度でおきているかを地の色の濃淡で表現している。これにより、ユーザは、その変異が実験上、避けられない塩基置換であるか、または、配列特有の特徴（例えば、その変異が、ある疾患の原因となる変異である場合など）であるか、実験によりその変異を元に戻さなければならないかの示唆が得ることができる。

【 0 0 8 8 】

また、相同性に基づいて対応する部分をそろえて配列を表示している。既存の

配列比較では、図 12 に示すように、入力された配列と比較する配列を個別に 1 対 1 で比較していた。これにより、ユーザは入力された配列と検索された配列で網羅的に変異の情報について把握することができ、実験操作によりその変異を元に戻すかの示唆を得ることができる。

【0089】

また、既存の配列比較結果である図 12 にリンクの情報がつけられている。また、ユーザが配列の何塩基目を見たいかを指定すると、リンクによりその部位における配列比較を表示することができる。これにより、ユーザは変異情報を網羅的に相互参照することが可能である。

【0090】

図 9 は、アミノ酸配列について相同性検索を行う際の入力画面を示している。ここでは、ユーザによって、アミノ酸配列の入力、検索を行うための DB の指定、P S I - B L A S T を行うかの指定、検索結果を限定するための生物種の指定、まだ解析の行われていない検索結果ファイルの指定が行える。

【0091】

ユーザがアミノ酸配列入力フォームにアミノ酸配列を入力した後、S u b m i t ボタンをクリックすると、クエリ解析部 11 がアミノ酸配列を受け取り、上述の各部の処理が行われる。

【0092】

本実施例では、P S I - B L A S T を用いて、検索結果を得るかどうかをユーザが指定できるようになっており、ここでの指定を、P S I - B L A S T フラグとして保持する。これにより、ユーザは検索結果を得る手段として、P S I - B L A S T を使うことができ、P S I - B L A S T でないと得られないような相同性をもつ配列について、比較することが可能である。

【0093】

また、結果を生物種によって絞り込むかどうかをユーザが指定できるようになっており、ここでの指定を生物種情報フラグに保持する。

【0094】

また、ユーザが従来の相同性検索によって得られた結果を解析対象として選択

することが可能である。

【0095】

図10は、アミノ酸配列について相同性検索を行う際の結果画面を示している。入力されたアミノ酸配列のうち、変異のある部位は赤い文字で表している。これにより、ユーザは入力した配列について変異部位についての全体的な状況を容易に把握することができる。

【0096】

また、比較した各々の配列は60字ごとに区切られている。

【0097】

また、相同性に基づいて対応する部分をそろえて配列を表示している。

【0098】

また、各々の変異において、それがどの程度であるかを地の色の濃淡で表現している。

【0099】

また、アミノ酸の挿入が見られた場合には、地の色を黄色にして表示している。

【0100】

アミノ酸は、生物について20種類が用いられているが、各々、疎水性度や酸性／塩基性度などの性質が決まっている。変異に伴って、これらの性質が変わると、配列全体としての性質が変わることが予想されるだけでなく、その変異箇所が酵素の反応中心の部位だった場合、酵素の機能が失われることとなる。それをふまえて、アミノ酸どうしの性質の違いを記述した変換行列を用いて、アミノ酸どうしの性質の近さを計算し、その値によって地の色の濃淡を変えて表現している。本実施例では、この変換行列として、図11に示すBLOSUM62なる変換行列を用いている。

【0101】

また、図10において既存の配列比較結果にリンクの情報がつけられている。また、ユーザが配列の何塩基目を見たいかを指定すると、リンクによりその部位を表示することができる。このことにより、ユーザは変異情報を網羅的に相互参

照することが可能である。

【0102】

図13を参照すると、本発明の第2の実施形態が示されている。本発明の第2の実施形態の相同性検索システムは、制御部15を有する点が第1の実施の形態と異なっている。よって、以下では制御部15に関する説明のみ行う。

【0103】

制御部15は、クエリ解析部11と、通信部12と、結果解析部13と、表示部14の動作を制御するものである。制御部15を有することにより、クエリを連続的に処理する、または、複数のDBを検索するなどの際、通信回線での混雑状況のモニター、または、解析結果の生成のモニターを行うなどによって、各部における処理のタイミングを制御することが可能となる。

【0104】

制御部15は、クエリ解析部11と、通信部12と、結果解析部13と、表示部14から処理状況の通知を受け、その通知に従って各部の処理が順次行われるように制御する。よって、第2の実施形態におけるクエリ解析部11と、通信部12と、結果解析部13と、表示部14は、各々の処理の終了時に制御部15に制御終了を通知する機能をさらに有し、制御部15からの動作制御に従って動作する。クエリ解析部11と、通信部12と、結果解析部13と、表示部14における処理状況の通知出力および動作制御入力、制御部の処理をコンピュータプログラムとして作成し、第1の実施形態に付加することによって、第2の実施形態を実現することができる。これにより、例えばクエリが多数あって自動的に検索および解析を行いたい場合でも、ユーザがすべてのクエリを入力する必要がなくなり、ユーザの負担を大幅に削減することができる。

【0105】

【発明の効果】

以上説明したように、本発明によれば、次のような効果を有する。

【0106】

第1の効果は、塩基配列およびアミノ酸配列の相同性検索結果における変異の部位、頻度など重要な意味を持つ情報が、ユーザによる相互参照や再検索なしに

得られることである。その理由は、複数の類似した塩基配列またはアミノ酸配列を含む相同性検索結果を入力とし、類似した複数の塩基配列またはアミノ酸配列における変異性と共通性に応じた視覚的な特徴を付加して表示情報として出力することができるからである。

【0 1 0 7】

第2の効果は、相同性検索結果である複数の配列における変異部分が一目でわかることである。その理由は、類似する複数の塩基配列またはアミノ酸配列における変異部位に関する視覚的特徴を有する表示情報を出力できるからである。

【0 1 0 8】

第3の効果は、相同性検索結果である複数の配列における共通部分が一目でわかることである。その理由は、類似する複数の塩基配列またはアミノ酸配列における共通部位に関する視覚的特徴を有する表示情報を出力できるからである。

【0 1 0 9】

第4の効果は、相同性検索結果である複数の配列全体とともに相同性の高い部位を表示できることである。その理由は、本発明の表示情報が、配列全体を含むことができるからである。

【0 1 1 0】

第5の効果は、マルチプルアライメントの結果における塩基位置のずれをなくして結果を表示できることである。その理由は、塩基およびアミノ酸の位置対応関係に従って整列された形式の表示情報を出力できるからである。

【0 1 1 1】

第6の効果は、比較を行いたい配列情報を予め収集しておくことや、所望とする配列情報が手元に無い場合に、配列と関連する情報、例えば、遺伝子名、アミノ酸名、疾患名、タンパク質名、等を用いて配列情報を取得しておくことができることである。その理由は、複数の生物情報DBに検索条件情報を送る手段を有するからである。

【0 1 1 2】

第7の効果は、相同性検索結果およびマルチプルアライメントの結果である配列において、変異部分および共通部分が一目でわかるような表示を行えることで

ある。その理由は、本発明が上記第 1 の効果から第 6 の効果を併せ持つことができるからである。

【0113】

第 8 の効果は、相同性検索結果を得た前後にユーザが行わなければならなかった作業を低減し、ユーザの負担を軽減することである。その理由は、マルチプルアライメントにおける位置合わせが不要であり、リンクを用いて相互参照も表示情報に包含可能であり、配列と関連する情報を用いて配列情報を取得可能であるからである。

【0114】

第 9 の効果は、検索条件を変えて何度も検索する場合でも、研究者が条件を変えては検索しなおすという作業を不要とし、効率的な相同性検索を可能にすることである。その理由は、検索の一連の手順を制御する制御部を有するからである。このような制御部を有することによって、公共の生物情報 DB へのクエリのタイミングを制御することもでき、公共の生物情報 DB サーバの負荷を軽減することにも有効である。

【図面の簡単な説明】

【図 1】

本発明の第 1 の実施形態の相同性検索システムの構成を示した図である。

【図 2】

クエリ解析部 11 において実施される処理の流れを説明した図である。

【図 3】

通信部 12 において実施される処理の流れを説明した図である。

【図 4】

結果解析部 13 において実施される処理の流れを説明した図である。

【図 5】

表示部 14 において実施される処理の流れを説明した図である。

【図 6】

塩基配列における相同性検索の入力画面の一例である。

【図 7】

塩基配列における相同性検索の結果画面の一例の一部である。

【図 8】

図 7 の塩基配列における相同性検索の結果画面の別の一部である。

【図 9】

アミノ酸配列における相同性検索の入力画面の一例である。

【図 1 0】

アミノ酸配列における相同性検索の結果画面の一例である。

【図 1 1】

アミノ酸置換行列 B L O S U M 6 2 である。

【図 1 2】

遺伝子のクエリ配列に対する従来の B L A S T 検索結果例である。

【図 1 3】

本発明の第 2 の実施形態の相同性検索システムの構成を示した図である。

【符号の説明】

1 0 相同性検索システム

1 1 クエリ解析部

1 2 通信部

1 3 結果解析部

1 4 表示部

1 5 制御部

2 1 サーバ

2 2 データベース

2 0 1 ～ 2 1 2 ステップ

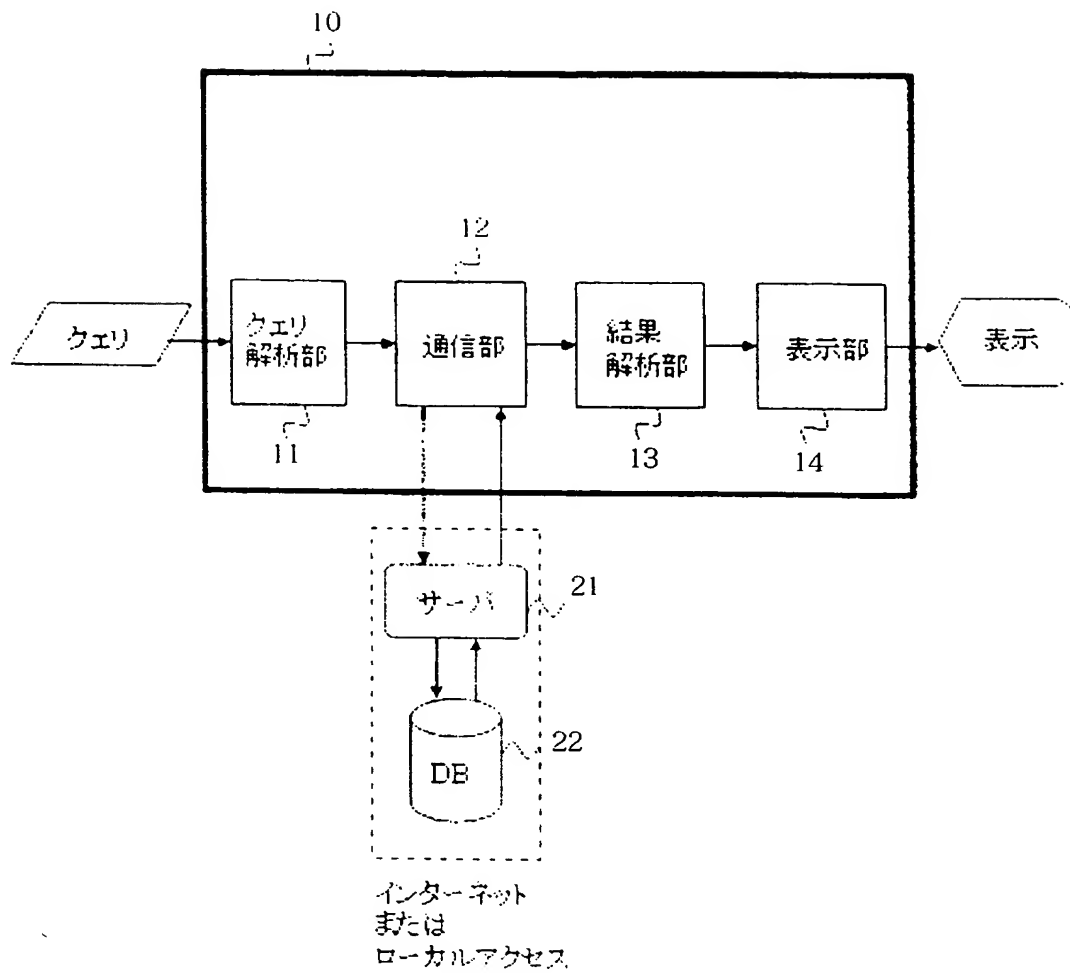
3 0 1 ～ 3 1 2 ステップ

4 0 1 ～ 4 0 5 ステップ

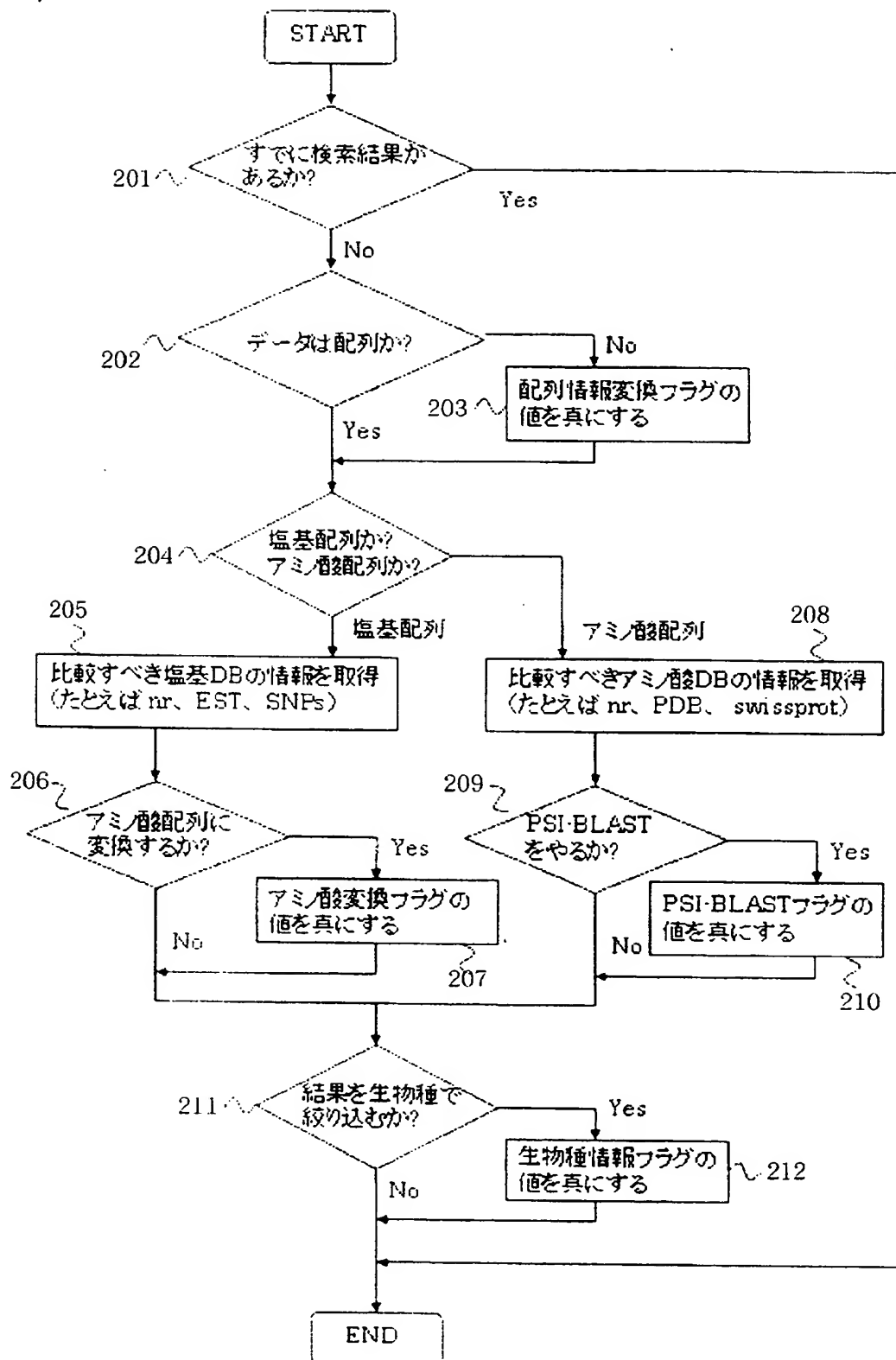
5 0 1 ～ 5 0 6 ステップ

【書類名】 図面

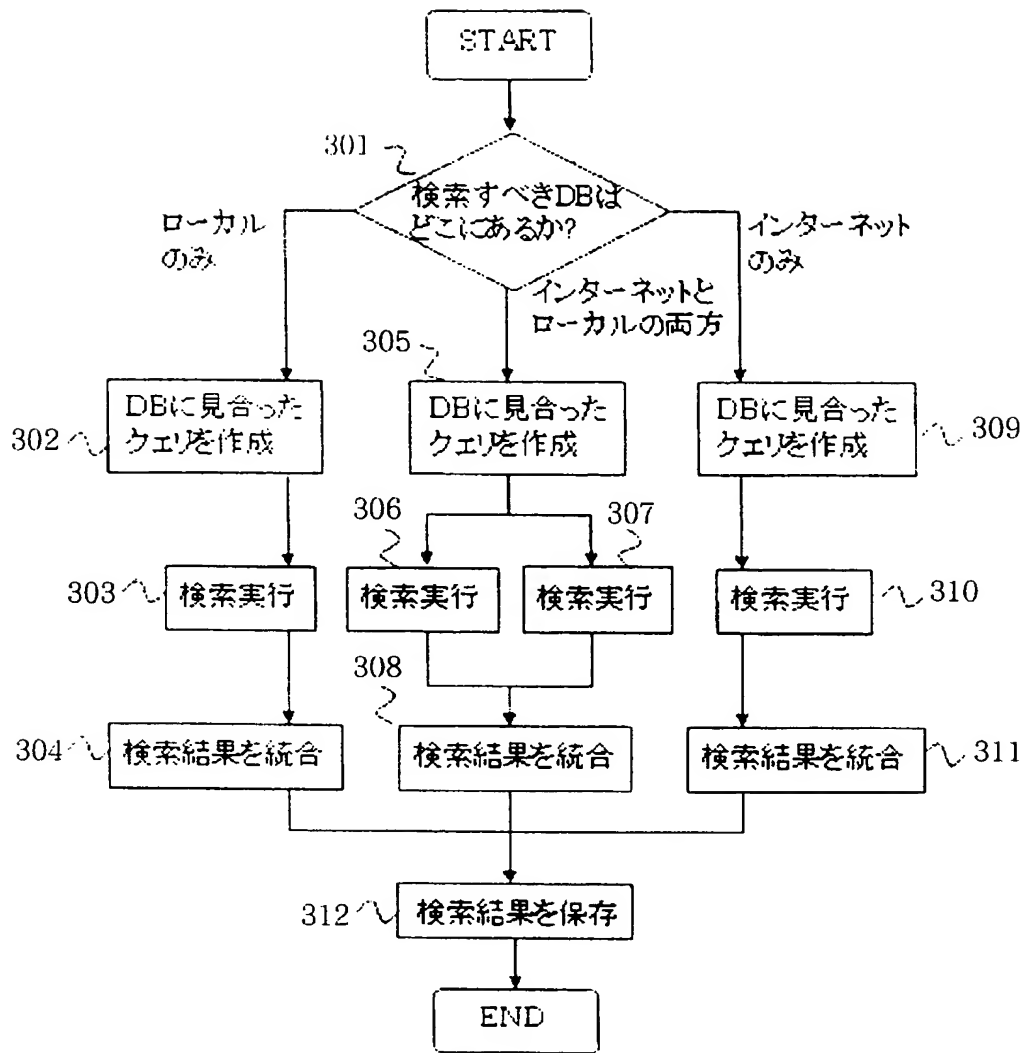
【図 1】



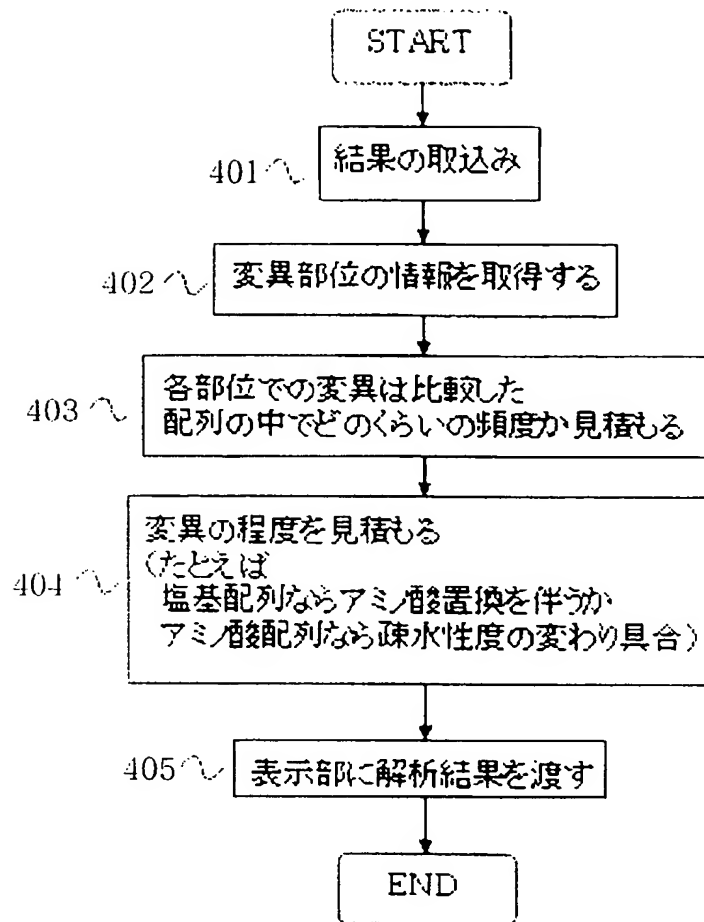
【図 2】



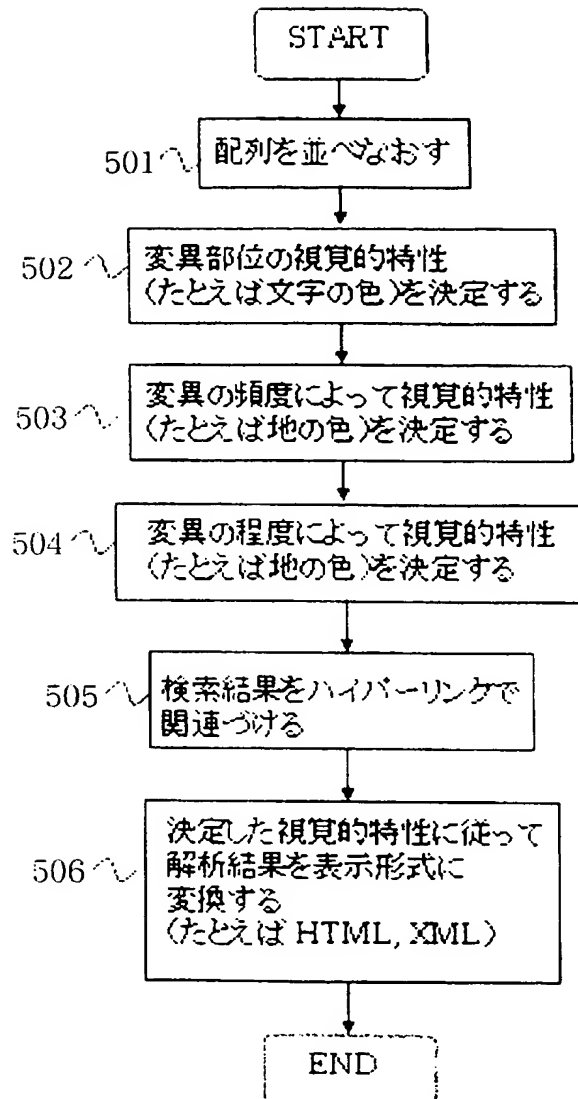
【図 3】



【図 4】



【図 5】



【図 6】

Blast Auto Submission Form

If you submit, your query send to NCBI BLAST, and parse results.

Enter your Query

or Enter Query File

Choose database ☒ Human nr ☒ Human EST ☒ Human SNPs

Limit by entrez query select from:

Display translated codon ☒ check this out, ex. query is promoter

To submit the query, To clear the form

Blast Result Parser Form

If you submit, parse BLAST Result.

Enter your Query

or Enter Query File

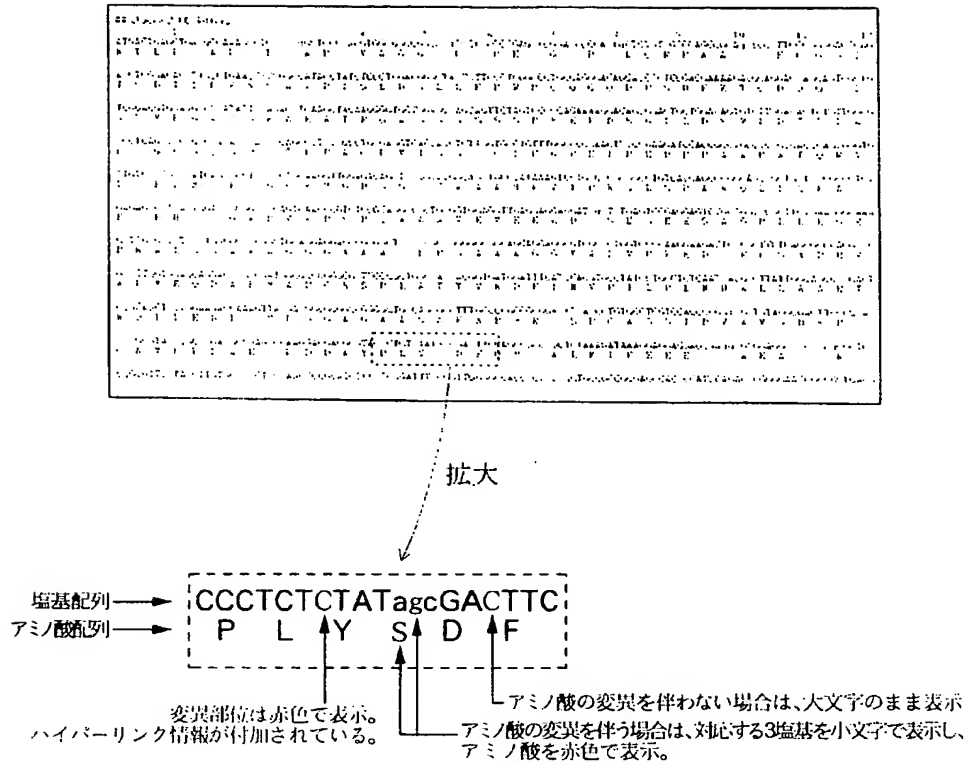
Enter Blast Result File

Enter Blast Result File

Enter Blast Result File

To submit the query, To clear the form

【図7】



ALIGNMENTS

```
Sequence ID >>
Query 1 - 60
gn1|dbSNP|rs2008112
g1|35651|emb|X51730.1|KSPREC
g1|4102792|gb|AF016301.2|AF016301
g1|4505766|ref|INV
g1|189934|gb|U15716.1|HUMPGRR
g1|22759851|dbj|AB085845.1|
g1|22759949|dbj|AB085844.1|
g1|22036116|dbj|AB085603.1|
g1|20278870|dbj|AB084248.1|
g1|22759947|dbj|AB085843.1|
g1|14245763|dbj|AP001573.4|
```

[illegible]

```

Sequence ID >>>
Query 1801 - 1860
g|15932498|c|eab|A1709143|1|A1709143|
g|15936808|gb|B1827058|1|B1827058|
g|15943587|gb|B1832037|1|B1832037|
g|22912092|gb|B0570792|1|B0570792|
g|24718342|gb|CA388684|1|CA388684|
g|24718342|gb|B0292212|1|B0292212|
g|14002748|gb|BC723061|1|BC723061|
g|11005321|dbj|A1438001|1|A1438001|
g|98650050|gb|A6744786|1|A6744786|
g|9865843|dbj|A6547483|1|A6547483|
g|9868947|dbj|A6744729|1|A6744729|
g|4687594|gb|A1636264|1|A1636264|
g|4222439|gb|A1393892|1|A1393892|
g|11002233|gb|B0446121|1|B0446121|
g|98665650|gb|A1439558|1|A1439558|
g|11001479|dbj|A1393958|1|A1393958|
g|5928014|c|eab|A1037033|1|A1037033|
g|2984958|gb|A0876157|1|A0876157|
g|9775354|gb|BES46709|1|BES46709|
g|18046606|gb|B312261|1|B312261|
g|15263665|gb|B1438915|1|B1438915|
g|2041283|gb|A9551606|1|A9551606|
g|13720136|gb|B017680|1|B017680|
g|13720136|gb|BF198449|1|BF198449|
g|11309168|gb|BF350094|1|BF350094|
g|3001851|gb|A0867643|1|A0867643|
g|10995564|dbj|A1138063|1|A1138063|

```

[illegible]

【図 9】

Blast Auto Submission Form

If you submit, your query send to NCBI BLAST, and parse results.

Enter your Query

or Enter Query File



Choose database

nr

Limit by entrez query select from: (none)

PSI BLAST iteration 1 means standard blastp

To submit the query,

To clear the form

Blast Result Parser Form

If you submit, parse BLAST Result.

Enter your Query

or Enter Query File



Enter Blast Result File



Enter Blast Result File



Enter Blast Result File



To submit the query,

To clear the form

Query 933 letters

[illegible]

ALIGNMENTS

Jump Cinar

Sequence ID 11

```
Query 1 - 60
gl 356521 cab CAA36018 1
gl 625331 pir QRPUR
gl 126441001 sp P06401 PRGR
gl 4505767 ref NP
gl 22759952 db BAC11013.1
gl 22759948 db BAC11011.1
gl 20278871 db BAB91074.1
gl 22036117 db BAC06505.1
gl 22759950 db BAC11012.1
```

[illegible]

Sequence ID 2: 88

Query 541 - 600

```

g1 3565231|par|CAA36010.1|
g2 6255331|par|QRHUP
g3 12644401|ref|NP_P064011|FRGR
g4 5055767|ref|NP
g5 22759952|db|BAC11013.1
g6 22759948|db|BAC11011.1
g7 22708071|db|BAB91074.1
g8 20361617|db|BAC0585.1
g9 22759950|db|BAC1102.1
g10 3342602|gb|AAC63513.1
g11 1788992|gb|AAA51774.1
g12 4504133|ref|KAP
g13 4586557|gb|AAA516603.1
g14 460281|gb|AAA51770.1
g15 4505195|ref|NP
g16 7217|par|QRHUGE

```

[illegible]

Sequence ID > < :

Query 601 - 660

```

g1 326552|amb|CAC56018.1
g1 3525331|pir|ORFPU
g1 12644100|ap|PG6401|PRGR
g1 4505767|ref|NF
g1 22759952|db|BAC11013.1
g1 22759948|db|BAC11011.1
g1 20278871|db|BAB91074.1
g1 22036117|db|BAC06585.1
g1 22759502|db|BAC11012.1
g1 3342402|gb|AAC63513.1
g1 178892|gb|AAAS1774.1
g1 4504133|ref|NF
g1 458657|gb|AAAB16603.1
g1 4602801|gb|AAAS1770.1
g1 4505199|ref|NF
g1 72117|pir|ORFHUGS

```

[illegible]

【図 11】

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V	B	Z	X	*
A	4	-1	-2	-2	0	-1	-1	0	-2	-1	-1	-1	-1	-2	-1	1	0	-3	-2	0	-2	-1	0	-4
R	-1	5	0	-2	-3	1	0	-2	0	-3	-2	2	-1	-3	-2	-1	-1	-3	-2	-3	-1	0	-1	-4
N	-2	0	6	1	-3	0	0	0	1	-3	-3	0	-2	-3	-2	1	0	-4	-2	-3	3	0	-1	-4
D	-2	-2	1	6	-3	0	2	-1	-1	-3	-4	-1	-3	-3	-1	0	-1	-4	-3	-3	4	1	-1	-4
C	0	-3	-3	-3	9	-3	-4	-3	-3	-1	-1	-3	-1	-2	-3	-1	-1	-2	-2	-1	-3	-3	-2	-4
Q	-1	1	0	0	-3	5	2	-2	0	-3	-2	1	0	-3	-1	0	-1	-2	-1	-2	0	3	-1	-4
E	-1	0	0	2	-4	2	5	-2	0	-3	-3	1	-2	-3	-1	0	-1	-3	-2	-2	1	4	-1	-4
G	0	-2	0	-1	-3	-2	-2	6	-2	-4	-4	-2	-3	-3	-2	0	-2	-2	-3	-3	-1	-2	-1	-4
H	-2	0	1	-1	-3	0	0	-2	8	-3	-3	-1	-2	-1	-2	-1	-2	-2	2	-3	0	0	-1	-4
I	-1	-3	-3	-3	-1	-3	-3	-4	-3	4	2	-3	1	0	-3	-2	-1	-3	-1	3	-3	-3	-1	-4
L	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4	-2	2	0	-3	-2	-1	-2	-1	1	-4	-3	-1	-4
K	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5	-1	-3	-1	0	-1	-3	-2	-2	0	1	-1	-4
M	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5	0	-2	-1	-1	-1	-1	1	-3	-1	-1	-4
F	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6	-4	-2	-2	1	3	-1	-3	-3	-1	-4
P	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7	-1	-1	-4	-3	-2	-2	-1	-2	-4
S	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4	1	-3	-2	-2	0	0	0	-4
T	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5	-2	-2	0	-1	-1	0	-4
W	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11	2	-3	-4	-3	-2	-4
Y	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7	-1	-3	-2	-1	-4
V	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4	-3	-2	-1	-4
B	-2	-1	3	4	-3	0	1	-1	0	-3	-4	0	-3	-3	-2	0	-1	-4	-3	-3	4	1	-1	-4
Z	-1	0	0	1	-3	3	4	-2	0	-3	-3	1	-1	-3	-1	0	-1	-3	-2	-2	1	4	-1	-4
X	0	-1	-1	-1	-2	-1	-1	-1	-1	-1	-1	-1	-1	-1	-2	0	0	-2	-1	-1	-1	-1	-1	-4
*	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	1

【図 12】

□>gi|20810385|cb|BC028258.1| **L****U** Mus musculus, solute carrier family 26, member 6, clone MGC:25824
IMAGE:4165725, mRNA, complete cds
Length = 2681

Score = 44.1 bits (22), Expect = 1.1
Identities = 85/106 (80%)
Strand = Plus / Plus

Query: 1282 gacagtaaccaggagctcgtgctctcggctcagtaacacagtgagggtctctccag 1341
|||||
Sbjct: 1220 gacagtaaccaggagctcgtagcccttgccctcagtaacctcattgagggtctctccag 1279

Query: 1342 tgctactgcgtcacctcctccatgtccctagccctggtcaggaga 1387
|||||
Sbjct: 1280 tgcttcccttgagctcctccatgtctcggagcttggtacaggaga 1325

Score = 44.1 bits (22), Expect = 1.1
Identities = 25/26 (96%)
Strand = Plus / Plus

Query: 1805 tgtactttgccaatgctgagctctac 1830
|||||
Sbjct: 1743 tgtacttgcgaatgctgagctctac 1768

□>gi|22775306|gb|AY049076.1| **L****U** Mus musculus anion exchanger SLC26A6a mRNA, complete cds
Length = 2552

Score = 44.1 bits (22), Expect = 1.1
Identities = 85/106 (80%)
Strand = Plus / Plus

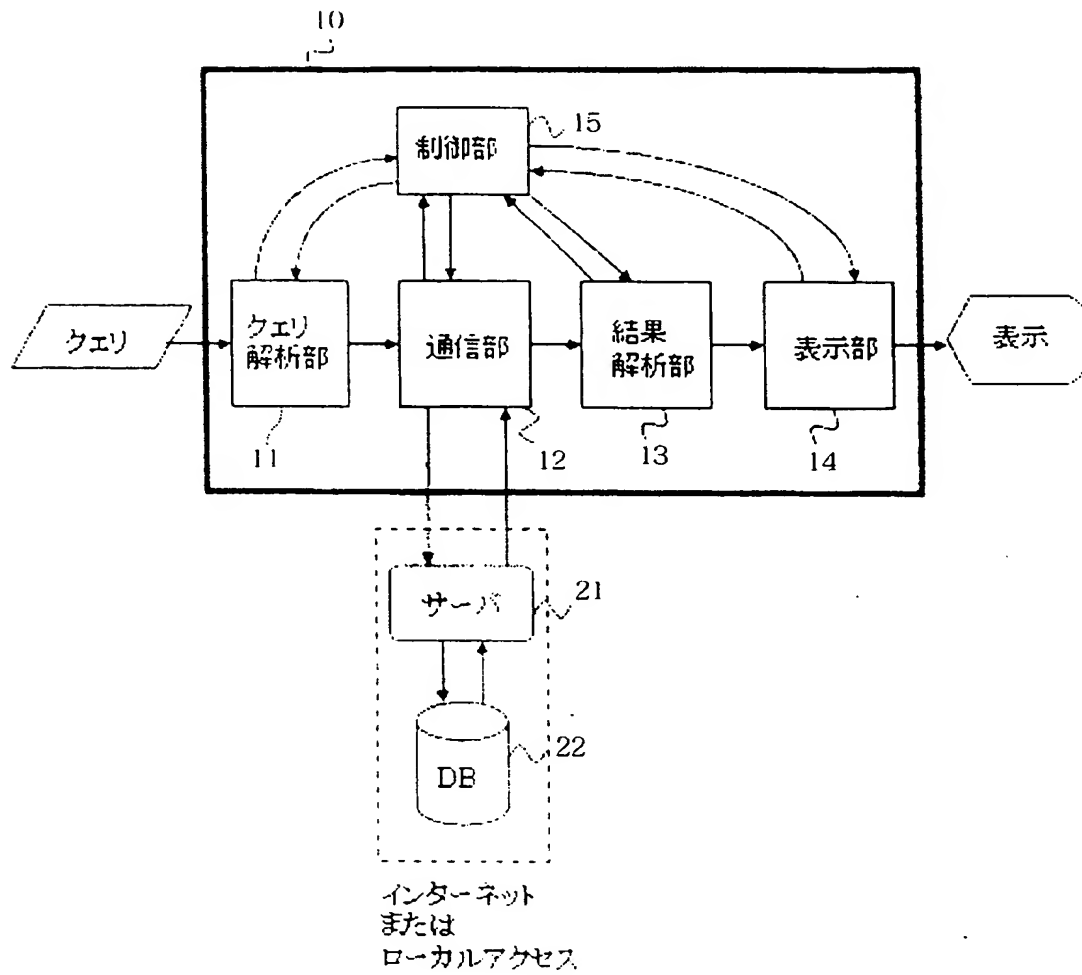
Query: 1282 gacagtaaccaggagctcgtgctctcggctcagtaacacagtgagggtctctccag 1341
|||||
Sbjct: 1156 gacagtaaccaggagcttgtagcccttgccctcagtaacctcattgagggtctctccag 1215

Query: 1342 tgctactgcgtcacctcctccatgtccctagccctggtcaggaga 1387
|||||
Sbjct: 1216 tgcttcccttgagctcctccatgtctcggagcttggtacaggaga 1261

Score = 44.1 bits (22), Expect = 1.1
Identities = 25/26 (96%)
Strand = Plus / Plus

Query: 1805 tgtactttgccaatgctgagctctac 1830
|||||
Sbjct: 1679 tgtacttgcgaatgctgagctctac 1704

【図 13】



【書類名】 要約書

【要約】

【課題】 配列間の変位部位や共通部位などに関する情報へのアクセスが容易で、検索された複数の配列をそのまま結果表示できる配列表示方法および相同性検索方法を提供する。

【解決手段】 本発明の配列表示方法は、類似した複数の塩基配列またはアミノ酸配列における変異性および／または共通性に応じた視覚的特徴を付加する。視覚的特徴は、変異部位および／または共通部位に付加され、変異性および／または共通性の程度、変異部位における変異の頻度に応じて付加される。さらに、視覚的特徴は、塩基配列をアミノ酸配列に変換し、変異部位のコドンに対応するアミノ酸名および性質を含むアミノ酸情報に基づいて付加される。類似した複数の塩基配列またはアミノ酸配列に関する情報へのリンクも設けられる。本発明の相同性検索方法は、この配列表示方法を用いて検索結果を表示する。

【選択図】 図 7

特願 2002-358407

出 願 人 履 歴 情 報

識別番号

[000004237]

1. 変更年月日

1990年 8月29日

[変更理由]

新規登録

住 所

東京都港区芝五丁目7番1号

氏 名

日本電気株式会社